When aspiration meets voicing: Native Danish speakers' production of Spanish initial stop voicing

Rane S. Ahle[1], Ocke-Schwen Bohn[2]
[1] Department of Spanish, Aarhus University, Aarhus, Denmark
[2] Department of English, Aarhus University, Aarhus, Denmark
ranesch@gmail.com, ocke.bohn@cc.au.dk

Abstract

The present study examines how native speakers of a language with a short-lag vs long-lag aspiration contrast (Danish) for initial stop consonants realize the voicing contrast (prevoiced vs short-lag) in their foreign language (L2), Spanish. We report and compare the voice onset time (VOT) of Danish and of Spanish initial stops as produced by 23 native (L1) Danish speakers who are learning Spanish as an L2 in high school. Both beginning and intermediate learners produced Spanish voiceless stops with significantly shorter VOTs than their Danish counterparts, but they never prevoiced, producing Spanish voiced stops with VOTs that did not differ from their L1 short-lag stops. These results fit in with previous studies which have shown that both experienced and inexperienced learners adjust, but do not reach, the VOT of L2 stops in the positive (aspirated) range. Our participants thus exhibit sensitivity to aspiration differences, but are apparently insensitive to prevoicing.

Introduction

As illustrated schematically in Figure 1, languages with a binary phonological voicing contrast for initial stop consonants typically implement this contrast using either aspiration (short-lag vs long-lag) or voicing (prevoicing vs short-lag).

The second language (L2) learning scenarios that arise from the difference between "aspiration" languages and "voicing" languages have been studied

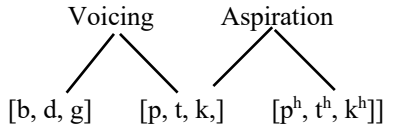

Figure 1: Typical implementation of the binary phonological voicing contrast through presence vs absence of voicing (left) or presence vs absence of aspiration (right).

mostly with respect to how adult L1 speakers of a voicing language with [p, t, k] produce and perceive the long-lag stops [p^h, t^h, k^h] of an aspiration language. These studies typically report compromise values for long-lag stops, i.e., values that are intermediate between the L1 short-lag and the L2 long-lag stops (e.g., Flege & Eefting, 1987). Even the highly experienced L1 Spanish speakers in the Garibaldi & Bohn (2017) study did not produce Danish long-lag stops with target values with but with compromise values.

A smaller number of studies has examined the reverse scenario, i.e., how adult speakers of an aspiration language produce and perceive the short-lag stops of a voicing language. These studies reveal the same pattern as the one mentioned above: An approximation of the target language VOT with compromise values, without actually reaching these values, both for relatively inexperienced and for experienced L2 learners. (e.g., Flege 1987, Major 1990).

To the best of our knowledge, no study has yet examined how L1 speakers of an aspiration language with short-lag [p, t, k] produce and perceive the prevoiced stops [b, d, g] of a voicing

language. Addressing this lacuna presents one of the motivations for the present study, which is based on Ahle (2023), and which asks the following questions:

- How do L1 speakers of an aspiration language produce the stops of a voicing language?
- Do learners show a sensitivity towards the L2 way of implementing the voicing contrast by either producing reduced aspiration intervals for [p, t, k] and/or prevoicing for [b, d, g]?
- Is the L2 production of the voicing contrast affected by the length of Spanish language instruction?

Characteristics of the two languages

Both the Spanish (SP) and the Danish DK) way using VOT for implementing the phonological voicing contrast for stops in initial position has been studied extensively. Even though these studies may differ widely in how they elicited tokens (e.g., following stressed or unstressed vowel, carrier phrase or not, etc.) the overall picture is clear:

SP has a contrast between initially prevoiced stops (range: ca. -90 ms to -60 ms VOT) and unaspirated stops (range: ca. 14 ms to 30 ms VOT), with increasing VOT values from anterior to posterior places of articulation.

DK has a contrast between initially unaspirated stops (range: ca. 6 ms to 35 ms VOT) and aspirated stops (range: ca. 30 ms to 140 ms). DK is remarkable in two respects: Unlike other languages with an unaspirated vs aspirated voicing contrast (e.g., English, see Dmitrieva et al. 2015), DK stops in initial stressed position are never optionally prevoiced. Also, whereas VOT increases in many (most?) languages as the place of articulation moves from anterior to posterior (Lisker & Abramson, 1964; Cho & Ladefoged, 1999), DK aspirated coronal stops typically have a longer VOT than velar stops (e.g., Fischer-Jørgensen 1954, Mortensen & Tøndering 2013, but see Puggaard 2023), and they are also heavily affricated, which is why they are transcribed here as [ts] following Schachtenhaufen (2022).

Methods

Participants

Two groups of adolescent L1 Danish speakers (age range: 16-19 years) participated. They were learning Spanish as a foreign language in high school in Aarhus, Denmark. Group 1g consisted of 13 students who had taken Spanish classes for 9 months with lessons of two times 90 minutes per week at the time of the study, and group 3g consisted of 11 students who had taken Spanish classes for more than two years with lessons of three times 90 minutes per week at the time of the study. The teachers for both groups were L1 Danish speakers, and the languages of instruction were both Danish and Spanish, with increased use of Spanish in the advanced 3g classes.

Materials and Procedure

Participants read Spanish and Danish disyllabic real words embedded in Spanish and Danish carrier sentences. The initial segment of the word was [b, d, g, p, t, k] for the Spanish tokens and [p, t, k, p^h, ts, k^h] for the Danish tokens. Because VOT is affected by stress (Lisker & Abramson, 1967) and by the following vowel (Port & Rotunno, 1979; Mortensen & Tøndering 2013), both the Spanish and Danish words had initial stress, they occurred in initial position in the carrier sentences, and the following vowel was either [i, a] or [u], which Danish and Spanish share.

Tokens were elicited by presenting individual power point slides for each word. Half of the participants read a randomized Spanish list first, and the other half, a randomized Danish list. Recordings were conducted individually in a quiet environment using Zoom H2n digital recorder. Each participant

recorded 72 tokens, resulting in (23 participants x 72 tokens) = 1656 tokens. A total of 189 tokens were not used for further analysis (because of technical problems, noise, or failure of participants to follow instructions). In addition, three prevoiced tokens, all from one participant's total of 72 tokens, were considered outliers and not included in the analysis. This resulted in a total of 1464 tokens for analysis.

VOT was measured using *praat* (Boersma & Weenink, 2022) following standard procedures, i.e., by measuring the interval between the onset of the release burst and the onset of periodicity in waveform displays.

Results

VOT in Danish initial stops

We first compared the mean VOT values produced by 1g and 3g for Danish stops using t-tests for each place of articulation. Table 1 shows the mean VOT values for the DK unaspirated stops. The two groups did not differ significantly ($p > .3$) with respect to any of the VOT values.

Table 1. Mean VOT values (in ms) as produced for Danish unaspirated initial stops by two groups of Danish high school students. SD in parentheses.

	p	t	k
1g	20.5 (8.6)	22.4 (10.5)	31.0 (11.8)
3g	20.5 (7.6)	24.9 (10.1)	32.2 (9.0)

The values for the aspirated stops are shown in Table 2. The mean VOT values for [ts] and [k^h] did not differ significantly between the two groups. However, the VOT values for [p^h] produced by 1g were significantly shorter than those produced by 3g ($p = .0289$).

Table 2. Mean VOT values (in ms) as produced for Danish aspirated initial stops by two groups of Danish high school students. SD in parentheses.

	p^h	ts	k^h
1g	64.8 (17.8)	107.1 (26.1)	88.9 (19.0)
3g	72.0 (19.8)	102.9 (26.1)	94.3 (24.0)

VOT in Spanish initial stops

Table 3 shows the mean VOT values produced by 1g and 3g for target Spanish prevoiced stops, and Table 4 the VOT values produced by 1g and 3g for target Spanish unaspirated stops. Separate t-tests for each place of articulation revealed that the differences between the two groups were nonsignificant ($p > .4$).

Table 3. Mean VOT values (in ms) as produced for target Spanish prevoiced initial stops by two groups of Danish high school students.SD in parentheses.

	b	d	g
1g	20.0 (7.6)	25.5 (8.5)	29.6 (9.1)
3g	20.0 (7.2)	26.7 (8.7)	30.7 (9.6)

Table 4. Mean VOT values (in ms) as produced for target Spanish unaspirated initial stops by two groups of Danish high school students. SD in parentheses.

	p	t	k
1g	48.6 (21.9)	63.8 (40.1)	60.8 (22.3)
3g	57.6 (24.2)	61.6 (27.2)	67.1 (67.1)

Comparison of Danish and Spanish stop productions

This section compares the production of phonologically voiced and voiceless stops separately for each group and place of articulation. We used paired t-tests to establish whether the participants

produced differences between SP and DK cognates. The figures below are box plots in which the vertical line within each box represents the median, the lower and upper lines of the box the lower and the upper quartile, respectively, and the whiskers extending from the boxes represent the range.

Phonologically voiced stops

Figure 2 and Figure 3 illustrate how the participants from 1g and from 3g produced the phonologically voiced stops in their L1, DK, and in their L2, SP. The figures show that the participants produced the SP target prevoiced stops as unaspirated stops. The paired t-tests revealed that the VOT values for the DK unaspirated and the target SP voiced stops did not differ significantly ($p > .4$) for 1g and for 3g for any of the places of articulation, except for the coronal stops produced by 1g, whose target SP voiced stops had significantly longer mean VOTs (25.4 ms) than their DK unaspirated stops (22.4 ms), $t(262) = 2.826$, $p = .00634$. These results indicate that both L1 DK groups produced stops which in the L2 have lead VOT as unaspirated stops whose VOT values were more or less the same as those of the L1 unaspirated stops.

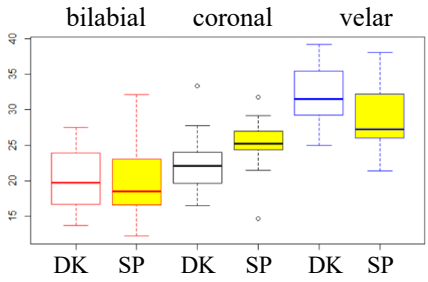

Figure 2. VOT values for DK unaspirated and SP target voiced stops as produced by 1g.

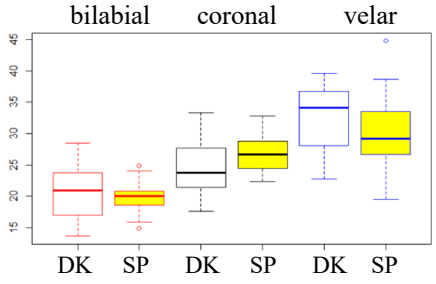

Figure 3. VOT values for DK unaspirated and SP target voiced stops as produced by 3g.

Phonologically voiceless stops

Figure 4 and Figure 5 illustrate how the participants from 1g and from 3g produced the phonologically voiceless stops in their L1, DK, and in their L2, SP. The figures show that both groups produced the SP target unaspirated stops as aspirated stops, but the figures also suggest that both groups produced these stops with reduced VOT values vis-à-vis their native DK values. This impression is confirmed by paired t-tests, which, for both groups and for all three places of articulation, revealed highly significant differences ($p < .001$) between the VOT values produced for SP and DK words.

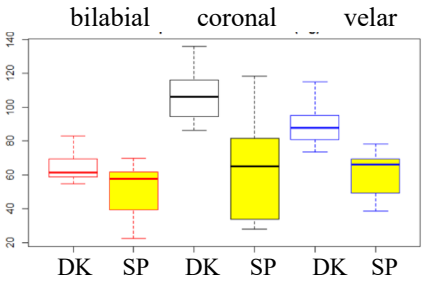

Figure 4. VOT values for DK aspirated and SP target unaspirated stops as produced by 1g.

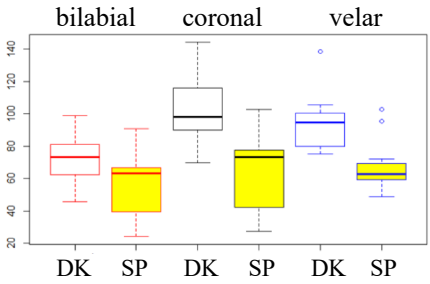

Figure 5. VOT values for DK aspirated and SP target unaspirated stops as produced by 3g.

Discussion and Conclusion

The present study examined how speakers of a language which uses aspiration to differentiate voicing in initial stops produce a voicing contrast in a language which uses voicing to differentiate initial stops. Two groups of L1 Danish speakers differing in Spanish language experience produced word-initial stops in stressed syllables for all three places of articulation in their L1, Danish, and in their L2, Spanish.

We found that both groups produced target unaspirated Spanish stops as aspirated stops with long-lag VOT values. However, these values were consistently and significantly shorter than the VOT values for the L1 DK aspirated stops, suggesting that the DK learners had developed a sensitivity towards the reduced VOT values for SP voiceless stops. Several earlier studies have likewise reported that even minimal exposure to a language which differentiates stops using voicing (prevoiced vs unaspirated) may result in the production of "compromise" VOT values by L1 speakers of an aspiration language, i.e., VOT values that are consistently intermediate between the values produced by native speakers of the languages (e.g., Flege & Hammond 1982, Flege & Eefting 1987).

No such sensitivity to the L2 values was found for Spanish prevoiced stops. We found that both groups produced target prevoiced Spanish stops as unaspirated stops whose VOT values did not differ from the L1 unaspirated stops.

There are two aspects about then present study that are remarkable: First, that the two-year difference in Spanish-language experience in high school classes between groups 1g and 3g had no effect on the production of VOT values for syllable-initial Spanish stops. Both groups produced compromise VOT values for Spanish unaspirated stops, and both groups did appeared insensitive to Spanish prevoicing, producing initial short-lag stops instead of target prevoiced stops.

It would be interesting to see in further studies how much Spanish-language experience (both in terms of quality and quantity) is needed for learners whose L1 has an aspiration contrast to show some sensitivity prevoicing in the L2 so that they will produce prevoiced stops. It would also be interesting to investigate whether this amount of experience would be different for L1 speakers of an aspiration language which does not have optional prevoicing (e.g., Danish) as opposed to an aspiration b language with optional prevoicing (e.g., English).

References

Ahle, R. S. (2023). Onkler eller Gud? En undersøgelse af dansktalende gymnasieelevers udtale af spanske plosiver. MA thesis, Aarhus University

Boersma, W., & Weenink, D. (2022). Praat: Doing phonetics by computer, version 6.2.23. [Software]. Retrieved from https://www.fon.hum.uva.nl/praat/, 17 October 2022.

Brown, E. K., & Copple, M. T. (2018). Constructing two phonological systems: A phonetic analysis of /p/, /t/, /k/ among early Spanish–English bilingual speakers. *International Journal of Bilingualism*, 22(1), 51-68.

Dmitrieva, O., Llanos, F., Shultz, A. A., & Francis, A. L. (2015). Phonological status, not voice onset time, determines the acoustic realization of onset f0 as a secondary voicing cue in Spanish and

English. *Journal of Phonetics*, 49, 77-95.

Fischer-Jørgensen, E. (1954). Acoustic analysis of stop consonants. *Miscellanea Phonetica 2,* 42-59.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15(1), 47-65.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, 15(1), 67-83.

Flege, J. E., & Hammond, R. M. (1982). Mimicry of non-distinctive phonetic differences between language varieties. *Studies in Second Language Acquisition*, 5(1), 1-17.

Garibaldi, C. L., & Bohn, O.-S. (2015). Phonetic similarity predicts ultimate attainment quite well: The case of Danish /i, y, u/ and /d, t/ for native speakers of English and of Spanish. In *Proceedings of the 18th International Congress of Phonetic Sciences*, 0245.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20(3), 384-422.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. Language and speech, 10(1), 1-28.

Major, R. C. (1990). L2 acquisition, L1 loss, and the critical period hypothesis. In Leather, J. & James, A. (eds.). *New Sounds 1990*, 14-25

Mortensen, J., & Tøndering, J. (2013, June). The effect of vowel height on voice onset time in stop consonants in CV sequences in spontaneous Danish. In *Proceedings of Fonetik* (pp. 49-52). Linköping University.

Port, R. F., & Rotunno, R. (1979). Relation between voice-onset time and vowel duration. *Journal of the Acoustical Society of America*, 66(3), 654-662.

Puggaard-Rode, R. (2023). Stop! Hey, what's that sound? Netherlands Graduate School of Linguistics. https://www.lotpublications.nl/Documents/631_fulltext.pdf

Schachtenhaufen, R. (2022). Ny dansk fonetik. BoD–Books on Demand.